DOCUMENT RESUME

ED 207 567 IR 009 679

AUTHOR Duncan, E. B.; And Others
TITLE Qualified Citation Indexing: Its Relevance to Educational Technology.
INSTITUTION Aberdeen Univ. (Scotland). Univ. Teaching Centre.; Robert Gordon's Inst. of Technology, Aberdeen (Scotland).
SPONS AGENCY Scottish Inst. of Adult Education, Edinburgh.
PUB DATE Apr 81
NOTE 11p.; Paper presented at the Educational Technology International Conference, (Aberdeen, Scotland, April 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Authors; *Citation Indexes; Costs; *Databases; *Educational Technology; *Indexing; *Information Retrieval; Online Systems; Relevance (Information Retrieval)
IDENTIFIERS *Qualified Citation Indexing

ABSTRACT

Citation indexing, which matches linked articles through links with authors rather than through subject keyword matching, is particularly relevant to educational technology, a widely-spread subject with a special user group of varying interests, difficult to cover in one retrieval service, and whose terminology is often ambiguous. By including links from lists of references, very large databases are created, some of whose links may be misleading. Qualified citation indexing seeks to refine the output by including terms that are mutually exclusive and unambiguous to describe the context of the reference. The Scottish Education Department Qualified Citation Indexing Project is setting up a citation database with linked references from citing to cited work, the links qualified by using a list of relational or descriptive terms compiled from previous studies and the suggestions of users. Retrieval will be tested and modifications will be built in from feedback thus acquired and a prototype index will be presented. A major part of the work will be the identification of key authors and key papers and estimates will be made of both computing and indexing costs. (Author/RAA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED207567

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Qualified Citation Indexing : its relevance to educational technology **

E. B. Duncan, D. Anderson and R. McAleese

Robert Gordon's Institute of Technology,
St. Andrew Street,
Aberdeen,
AB1 1HG.

University Teaching Centre,
University of Aberdeen,
Aberdeen,
AB9 2UB.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

E.B. Duncan

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

** A project supported by the Scottish Education Department entitled Qualified Citation Indexing 1980 to 1982

IR009679

Educational technology as a discipline or broad subject area, displays several features which form classic issues for discussion among information specialists. As a subject, it has no clear boundaries. The spread of interest crosses the path of many other specialist subject areas, thus making it very difficult to define from the point of view of information retrieval. It is a subject area which comprises a highly practical element at the same time as enveloping an element of very complex research and highly theoretical work. The merging of these two aspects has already been seen through the literature to cause severe communication problems, and consequently problems of information transfer.(CLARKE & DUNN, 1977; ELY,1979; GILBERT, 1977; HILLS & GILBERT,1977; McALEESE,1978) It is an area within which fashions can change very quickly - what is fashionable to study today may be tomorrow's taboo. For example, television in the nineteen seventies is now superseded by microcomputers. Unless one keeps up with the constantly changing emphasis, the terminology used in retrieving information may be totally unsuitable.

A second important feature is that while the subject spread is wide, it is probable that the users of educational technology information, may be a select few. For example, AETT has only 268 individual members. The user body is difficult to identify, as it may be composed of people with very different subject backgrounds and interests. For an information system to work effectively, it must be in close contact with the needs of people who are trying to obtain information from it. This involves the subject coverage of the base, but also to a very large extent the terminology used, and the inference of terms like 'information' for example in different contexts and to several users will mean something quite different. A retrieval system must therefore attempt to identify the user population and acquire as much feedback from it as possible. That is, the system should 'interrogate' the user to determine the nature of the enquiry, and permit the user to'browse' in a number of ways. (FORD,1977; GARVEY,1970; BROADFOOT,1979; HOUNSELL,1980) This can be a formidable problem, particularly for the larger systems and for any claiming international coverage. Even within a body of educational technologists, which a Conference such as ETIC will attract, there are, one suspects, many levels of understanding of different topics, and many modes of application of apparently similar information. These considerations are quite independent of the more obvious one,perhaps, of foreign language translation problems, which apply to any subject area. One would not wish to introduce a note of despair, but educational technology does have these particular features of difficulty for an information system. The user body has been mentioned as being small and specialised. One can refine this statement to mean that the body of people using the literature of educational technology is small, but in a sense we are all, that is, anyone involved in teaching or communication of information at its simplest level, users of the products of educational technology and this introduces another element which is very quickly apparent to anyone coming in to the subject area: the degree of communication of research findings to practical situations is low. (CLARKE & DUNN,1977; GILBERT,1977) Many excellent research projects seem to find little support in practical application, and as corollary, ideas springing from difficulties met in practice seem to have little chance of support or follow-up by research for wider trial or distribution. This situation may well be a direct consequence of poor information retrieval systems, or of systems not reaching or getting through to the appropriate user population in the code or language of the user. There is very little evidence, for instance,

of any effective SDI (selective dissemination of information) systems in educational technology, by which relevant selected information in some packaged form can be circulated to specific users or to small groups of research workers by matching output with pre-determined subject profiles i.e. individualised information.

It would be a relatively simple task to draw up a list of sources which claim to cover literature on educational technology from, say the British Library's collection list (KIST - keyword index to serial titles). The latest encyclopaedia on educational technology (UNWIN & MCALEESE 1978) lists 38 'easily accessible and useful English-language journals', of which four are what would be termed secondary sources, i.e. bringing together titles or abstracts of primary articles in meaningful subject groups. This is omitting some of the more general sources which will include educational technology, such as ERIC, with the associated Research in Education and Current Journals in Education, and British Educational Index. Another general source,which may not be so familiar in this field, is Social Sciences Citation Index, or even Science Citation Index, both of which include much material relevant to educational technology, by virtue of their cross-disciplinary spanning. To explain further why this should be so, it may be worth spending a little time describing the principle on which citation indexing is based and its relevance to educational technology. (MARTYN, 1965; WEINSTOCK, 1971; HALL, 1970; MARTYN, 1975)

A secondary source dealing only with literature from educational technology will miss many large areas. As has been pointed out, in a subject with as wide a spread of interest, this may be the most vital aspect of the retrieval, to draw from the whole spectrum of application in varying subject fields.

Take a subject such as 'Examinations' (see Fig. 1). A traditional or hierarchical approach might be to say that examinations are part of the general process of Certification, which in turn is to do with the purpose of Assessment; this is one among others of the reasons for Assessment, others being to test curriculum, or competence perhaps; further, the purpose of Assessment is only part of the general subject of Assessment, which is also concerned with problems of assessment, design of tests etc. Another user, however, might say that Examinations is too broad a description of his area of interest, and might want to make subdivision of this into types of examination, as indicated in Fig. 1. One can see that an approach like this may be very helpful, particularly if the user has little knowledge of the subject area, in guiding a user through the way in which the material in the index has been arranged. But the difficulty lies in the fact that this is only one way of looking at this whole subject area. It is one person's or organization's conceptual map of the subject. If this corresponds with the user's map, this is helpful. If, on the other hand, the reason for wanting information about O-grade results was to see if, for example, class size was an important factor, or whether physics should be taught to third year pupils, individually or in groups, for a two-hour or four-hour period, before or after lunch, i.e. the user's conceptual map was of a sociological nature in this instance, then the hierarchical type of approach may not be helpful.

In an attempt to satisfy both these criteria, the approach of citation indexing is to assume that in many instances a user will have a specific problem in mind - they know that X in Newcastle did some work on it, and that in fact, there was a good article defining just what it was, but it was several years ago and is now out of date. Rather than thinking of trying to define it in terms of subject keywords or patterns to match against an existing index, the approach is to

concentrate on the people considered to be key writers or workers on the topic. The basic assumptions made are (1) that the key author will have written papers with lists of references, but, (2) and more importantly, that other authors or workers will have quoted this author in their lists of references. Citation indexing thus takes the lists of references, or cited publications, and indexes those in such a way as to link them with the authors citing or quoting them, thus working forward in time rather than back. This can be seen to be a subtle way of covering a very broad spread of linked subjects, since a key paper on O-grade results, might have been quoted by someone writing on achievement testing, someone looking at class sizes and also by someone interested in, say, social class and learning. By finding People who have quoted this paper, one can then follow up similar work in a meaningful way, without the restriction of a framework of keywords or subject hierarchies or language or time divisions. One can demonstrate this very clearly in educational technology by taking a definable topic of interest, such as 'evaluation'. An educational technologist, knowing a little about the field, might be able to say immediately that there have been one or two 'key' people working in that area - one paper which keeps being quoted, for instance, being that by Parlett and Hamilton, first published in 1972 as 'Evaluation and Illumination'. By following these two key authors through SSCI, it will be found that this paper has been quoted not only by other 'key' people working on evaluation, but by writers in librarianship, physics, computing, health, psychology and so on. These papers will all have been written since the 'key' paper, and one can choose approximately from the indication of journal title, which will be of interest. Thus one can follow a 'key' paper forward in time. One can, of course, still trace backwards as well, as each citing paper will quote others which may also be relevant and so on. In fact, the main disadvantage of citation indexing is that it is almost too helpful, and that like many other information systems, it is in danger of producing more information than a user can absorb. It does have the same disadvantage of time delay as traditional indexing, in that any secondary source, by definition, depends on first receiving and absorbing the current literature - it is possible, however, that the digestion process is shorter, as there is less emphasis on subject division and classification.

We consider that citation indexing, for the reasons given above, is particularly suitable for application to educational technology. The point at which our project starts, is to examine ways in which a citation index might be refined, and to assess, by feedback from users, the most meaningful ways in which this might be achieved. We are setting up a qualified citation index in educational technology, indicating not only links between papers and those who quote them, but also the kind of link, or kind of relationship between cited and citing papers. It is in this sense that the index will be qualified - the nature of the relationship between papers will be indicated. To do this, it is necessary to examine each reference in the context in which it appears and assess, according to some 'formula' the author's reason for citing or quoting this reference. We started with an arbitrary list of relationship terms or 'qualifiers', taken from several sources (Table 1). There have been several papers written describing systems previously attempted (LIPETZ, 1965; WEINSTOCK, 1971; VOOS, 1976; FROST, 1979; HUANG, 1968; MARGOLIS, 1967; OPPENHEIM, 1978), but terms were taken also from random suggestions made by people in the course of their own reading. A very small pilot study was undertaken with two educational technologists at this stage using selected articles and personal interviews and discussion of them, and although this did yield

more and very helpful suggestions, it was decided that it would be extremely time-consuming and expensive to do this on a larger scale at this stage. Instead, this arbitrary list of terms was distributed,with introductory comment, to a group of lecturers and to a class of post-graduate librarianship students, all of whom were assumed to be involved in reading to some extent! They were asked to mark any term on this list which described the relationship between references and text of articles read, and in addition, if the description did not fit, to try to describe the relationship in some other way. We conducted a similar pilot study with authors of papers. We therefore ended up with a kind of league table of relationships which were found to be useful, and additional possibilities. This is the type of very simple consumer test we wish to carry out in educational technology, with as many interested users as possible. We have used the 'top 20' terms on articles included in the data base constructed so far and will now test reactions to these from educational technology users. Ideally, the relationships to be effective must be unambiguous and clear, and must be mutually exclusive. They must also be sufficiently easy for an indexer to apply unequivocally, so that the total indexing time and therefore cost, is not increased unnecessarily. The effect of these terms or 'qualifiers' on the output from the data base will be that instead of being able to get an answer to the question 'Who has done similar work to A on this topic?', we will now be able to ask 'Has there been a review article written along the same lines as the paper by A?', or 'Who else has used A's paper in a historical sense?', or 'Can I find another 'key' paper on a topic similar to that investigated by A?'.

We are thus approaching what we regard as a semi-intelligent system. Our aim will be to produce something which will be recognisable as English on-line, which we can 'chat' to, or interrogate informally, or browse through, i.e. the user can be creative and change strategy on receiving further prompts, to select relevant information from the mounds with which we seem to be perpetually presented. Another facility which has been included in our records is a simple keyword coding. This is not to be regarded as a principal search term, but as a secondary help, if an author is not known. While one is naturally looking for information on a subject, one must remember that the subjects are in no way grouped or even rationalised to any great extent - as explained previously, there is intentionally no requirement from the user to fit his subject search in to a pre-determined pattern. The keywords are intended to allow a user to find out which are the 'key' authors or 'key' papers on a topic, and take the search by author from there. This area is again open to discussion and modification by interaction with users. Other modifications will almost certainly be made,in the same way,to the length of each record, for example, the abbreviations of titles, from consideration of user comment and over-all cost of input.

As has been demonstrated by ISI, many permutations of the information are possible, once a format is agreed, and specialist output, such as print-outs of review articles on a specified topic,would be relatively simple spin-off effects. There are also refinements which one can carry out on the base itself, such as building in automatic counting or 'weighting' of references, so that assessment of various publication patterns might be made from time to time.

We hope our findings will make a small but significant contribution to information transfer within educational technology.

★ We would like to thank a number of colleagues who have helped us in this project: in particular John Cowan, Nick Rushby and Bernie Stocks.

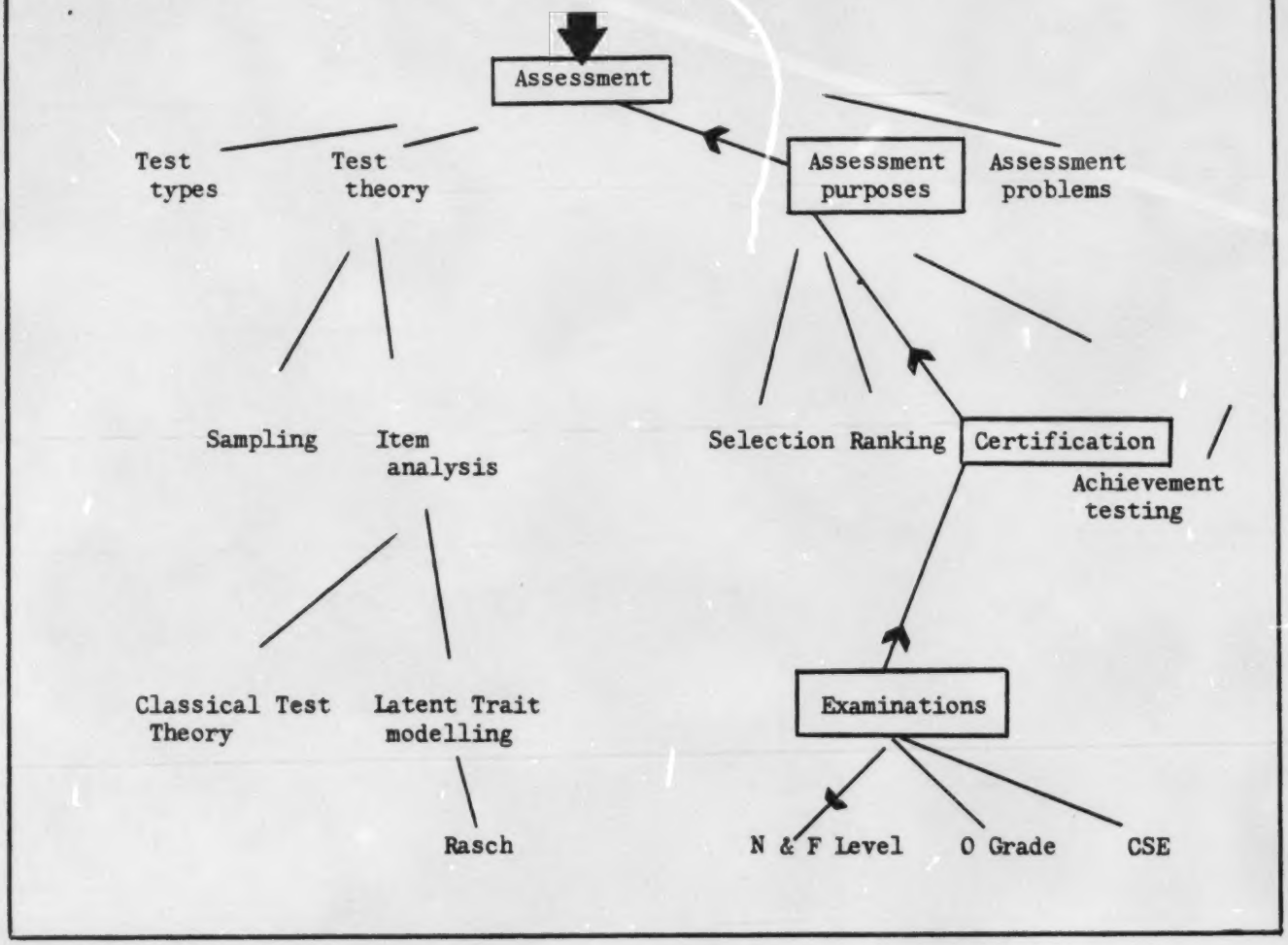

Figure 1: The relationship between Assessment, Examinations etc
(subject headings taken from the SED Information System restructured by A. Lyndsay for the SED)

- Paying homage
  Credit
- Bibliographical leads for researchers
  Review
- Historical
  Background reading
  Narrative
- Definition
  Clarification
  Illustration
- Anticipated value
  New research
- Methodology
  Equipment
- Correction
  Criticism
  Disputing
  Disclaiming
  Contradictory research
- Substantiation
  Corroboration
  Similar research
  Similar work
- Statistics
  Data
  Practical application

Table 1: Relational Qualifiers (i.e. descriptions of the type of relationship between a reference quoted by an author and the context of the quotation).

References
(with possible Relational Qualifiers)

	Relational Qualifiers
BRITISH LIBRARY. Keyword Index to Serial Titles (KIST)	Bibliographical leads
BROADFOOT Patricia (1979) Educational research through the looking glass. Scottish Educational Review 11(2) 133-142.	Corroboration
CLARKE Laurence J & DUNN William R (1977) Initial study of the use of research in educational technology by practitioners of educational technology and curriculum development. Final Report to the Council for Educational Technology. (spiral) 41pp.	Similar research
ELY Donald P (1979) The dynamics of an information system for educational technology. In-Educational technology twenty years on. Edited by G Terry PAGE & Quentin WHITLOCK. London, Kogan Page. (Aspects of Educational Technology. XIII) pp. 310-316.	Background reading
FORD Geoffrey (Editor) (1977) User studies: an introductory guide and select bibliography. University of Sheffield Centre for Research on User Studies. 92pp. (236ref.) (CRUS Occasional Paper no. 1)	Corroboration Review
FROST Carolyn O (1979) The use of citations in literary research: a preliminary classification of citation functions. Library Quarterly 49 399-414.	Similar research
GARVEY William D; Lin Nan; NELSON Carnot E (1970) Communication in the physical and social sciences. Science 170(3963) 1166-1173.	Corroboration
GILBERT J K (1977) An investigation into the use of research findings by some practitioners of educational technology in England. London, CET.	Background reading Similar research
HALL Angela M (1970) Use and value of citations: a state of the art report. London, Institution of Electrical Engineers. (INSPEC Report no. R70/4)	Definition

	Relational Qualifiers
HILLS Philip & GILBERT John (Editors) (1977) The spread of educational technology. London, Kogan Page. 474pp. (Aspects of Educational Technology. XI) p.263.	Background reading
HOUNSELL Dai; MARTIN Elaine; NEEDHAM Gill; JONES Haydn (1980) Educational information and the teacher. British Library, Boston Spa. (BLRD Report 5505)	Corroboration
HUANG Theodore S (1968) Efficacy of citation indexing in reference retrieval. Library Resources and Technical Services 12(4) 415-434.	Similar research
LINE Maurice B (1981) The structure of social science literature as shown by a large-scale citation analysis. Social Science Information Studies 1 67-87.	Clarification
LIPETZ Ben Ami (1965) Improvement of the selectivity of citation indexes to scientific literature through inclusion of citation relationship indicators. American Documentation 16 81-90.	Similar research
MCALEESE R (1980) The nature and dissemination of educational technology: lessons and implications. Paper presented at the 16th National Conference of the National Society for Performance and Instruction, San Francisco, 21-25 March 1978. Scientia Paedagogica Experimentalis 17(1) 44-60.	Background reading
MARGOLIS J (1967) Citation indexing and evaluation of scientific papers. Science 155 1213-1219.	Similar research
MARTYN John (1965) An examination of citation indexes. Aslib Proceedings 17(6) 184-196.	Definition
MARTYN John (1975) Citation analysis. Journal of Documentation 31(4) 290-297.	Definition
OPPENHEIM Charles & RENN Susan P (1978) Highly cited old papers and the reasons why they continue to be cited. Journal of the American Society of Information Science 29 225-231.	Similar research

	Relational Qualifiers
UNWIN Derick & MCALEESE R (1978) Encyclopaedia of educational media communications and technology. London, Macmillan.	Bibliographical leads
VOOS Henry & DAGAEV Katherine S (1976) Are all citations equal? Or, did we op.cit. your idem? Journal of Academic Librarianship 1(6) 19-21.	Similar research
WEINSTOCK Melvin (1971) Citation indexes. In-Encyclopaedia of librarianship and information science. Volume 5. Edited by Allen KENT & Harold LANCOUR. New York, Marcel Dekker. pp. 16-40.	Definition

.....